Quality Assurance and Quality Control Samples

Documentation includes:

- Complete documentation of methods used for all data collection, including explanation of any deviations from the prescribed methods;

- Explanation of all assumptions made in the data collection or analysis;

- All raw data, including identification of the source of each data point;

- All calculations, including copies of work conducted manually and all electronic spreadsheets or databases;

- Records of all relevant communication with team members and data contacts;

- QA/QC records, including responses to issues identified by *audits; and*

- Identification of sources of emissions not included in the inventory.

The source and type of the raw data will determine what type of information you need to place in the project file.

If the data were collected from...	Then you must maintain the following records...
Surveys	Original survey forms
Site visits	Site visit notes and reports
Source test reports	Complete copies of the reports
Internet pages or electronic bulletin boards	- Hard copy printouts of the pertinent data - Electronic copies of complete original data - Complete reference citation
Published document	- Complete reference citation - When possible, copies of the pages with the data used in the inventory
Unpublished documents or reports	- Complete reference citation - Copies of all pages with data used in the inventory - When possible, a copy of the entire document

If the data were collected from...	Then you must maintain the following records...
Personal communication - written	- Complete reference citation (contact name, affiliation, address or phone number, data of communication) - Copies of all pages with data used in the inventory - When possible, a copy of the entire document
Personal communication - verbal	Standardized Contact Report Form should be completed to record information obtained by telephone or at a meeting. An example Contact Report form is presented in EIIP Volume VI, Chapter 2.

You should include the following information on all manual calculation sheets:

- Preparer's name;

- Date prepared;

- Date modified, if applicable;

- Signature of reviewer;

- Date of review;

- Citations for all data used in the calculation;

- Explanation of all assumptions; and

- Page number on each page, showing total number of pages.

You must also document calculations performed using spreadsheets. The spreadsheet contains all of the pertinent information used to estimate emissions, but much of the information is often hidden on the summary printout. You must be certain that spreadsheets include the following information:

- Preparer's name;

- Date prepared;

- Date modified, if applicable;

- Spreadsheet version number;

- Name of reviewer;

- Date of review;

- All constants, factors, or other data used in the calculations must be presented; that is, all data needed to reproduce the value should be shown;

- Citations for all data used in the calculation;

- Explanation of all assumptions; and

- Page number on each page, showing total number of pages.

Refer to Appendix N for detailed instructions on procedures to follow when developing, documenting, and reviewing spreadsheet data.

3.13   What Documentation Procedures Must Be Followed?

For all steps in the inventory process, you should follow these procedures:

- Each member of the project team should be assigned a numbered, project-specific, notebook for recording all calculations and assumptions;

- All entries should be initialed and dated;

- Calculations and documentation should be done in ink - not pencil; and

- Errors should be corrected by drawing a single line through the original entry and writing the correct value nearby. All corrections must be initialed and dated by the person making the change.

In order to maintain credible inventory records, the inventory preparation plan will establish documented procedures for:

- File contents: Each member of the team should be aware of what materials must be placed in the project files;

- Location of all paper and electronic records;

- Ensuring that all hard copy records, as well as electronic databases and spreadsheets, are kept up-to-date;

- Adding data or background documentation to the file, including log-in procedures;

- Data tracking: To document who entered or manipulated data;

- Access to files: Checkout policies for paper documents, computer security policies for electronic files;

- Backup and maintenance of electronic files; and

- File name conventions.

3.14 How Do I Track Data Entry?

The procedures you use for tracking data entry will be established in the inventory preparation plan. Emissions inventory efforts often generate large quantities of data from a variety of sources. Development of emissions estimates generally requires data manipulation, including manual calculations, electronic data entry into computer programs, and data checking activities. Sometimes data will change hands several times during the emissions estimation process. This makes it very important that you use the procedures prescribed in the inventory preparation plan to record the names of the persons responsible and the dates of each data manipulation activity.

3.15 How Do I Check the Accuracy of Data Inputs and Manipulations?

The procedures you use for checking the accuracy of data entry and calculations will be established in the inventory preparation plan. It may be possible to build some of these checks into the computer database that is handling the data. For example, a computer program can be set to reject an entry that exceeds a reasonable or expected value. More often, manual QC checks, tracking the data from receipt to final result, are necessary.

It is important to check the transcription of data during the inventory preparation. Data manipulation steps that you need to monitor include:

- Transcription of data from raw data sheets into electronic spreadsheets or for use in manual calculations;

- Transcription of data results into summary tables or reports.

You must also have procedures in place to confirm the accuracy of calculated values. Many of these checks can be conducted with computer programs, others will require manual checks. Refer to Appendix N for guidance on evaluating accuracy of spreadsheet data.

NTI METHOD VERIFICATION SHEET

Team Member: ————————————— Source Category Name: ——————————

Brief Description of Category:————————————————————————

Specific NTI Source Category Code(s) (SCC, SIC, NTI CATCODE) If Applicable:————————
——

List of Useful References For:

Emission Factors: ——————————————————————————————
——
——

Activity Data:——————————————————————————————————
——
——

Methodology:——————————————————————————————————
——
——

Surrogate Activity Parameters Used for Spacial Apportioning (If Applicable):——————————
——
——

List of Additional References Reviewed But Not Useful:———————————————————
——
——

Methodology/Approach:————————————————————————————
——
——
——
——

Level of Detail Available (National/State/County):——————————————————————
——

Area Source Entry Form

NTI CATCODE: [ ]

AMS Code: [ ]

Pollutant ID: [ ]

Area Source Group ID: [ ]

Person Estimating: [ ]

Area Source Record Number: [example form]

Baseline Year: [ ]

If TRI data were used in estimate, provide year (eg 96 = 1996): [ ]

Data Entry Date: [ ]

QAQC Reviewer: [ ]

QAQC Date: [ ]

Methodology:

Calculations:

National Estimate: [0]

Units for National Estimate (tons/yr): [ ]

List of QA Documentation Forms Used to Incorporate ESD Template Data into the NTI

1. ESD Data Template Screening Sheet

The individuals performing the initial screening of the templates returned by ESD used the Screening Sheets to document which templates were returned with data, provide a brief summary of the data returned, and recommend the next step in the processing of the templates.

2. ESD Data Processing Summary Sheet

The individuals performing the duplicate record checks of ESD data against state data used these forms to document the actions taken to (1) identify duplicate facilities, (2) separate area and major source data, (3) apply the selection criteria to the duplicate facilities, (4) supplement the preferred data with any additional data , such as SIC, zipcode, etc., from the duplicate, and (5) perform any additional processing steps.

3. ESD-to-ORACLE Data Template Fact Sheet

These forms were used to summarize and quantify the tables of ESD data created after the processing step. Descriptions were provided for site, unit, process, stack, and related emissions data.

4. Processing QA Review Form

This is a QA documentation form for the manipulations performed which were documented in the ESD Data Processing Summary Sheet and the ESD-to-Oracle Data Template Fact Sheet.

5. Oracle Ready (Access) Tables: ESD Adds - Line Record Checks

This form documents that a QA was performed which ensures the Oracle-ready version of the ESD data matches the Access tables created as a result of the ESD data processing.

Table 1: Numeric Key to the Cut Criteria :

(1) omitted (discontinued)

(2) Are the emissions actual emissions? Actual emissions were given preference over potential or permitted emissions.

(3) Are the emissions TRI-based? Emissions estimates that are not TRI-based were given preference over TRI-based emissions records.

(4) Are the emissions at the SCC or process level? Emissions that were reported at the process level were given preference over emissions reported at the facility level.

(5) Are stack data included? Emissions records with stack data included were given preference over records without stack data.

(6) omitted (discontinued)

(7) ESD data is preferred. All other factors being equal, ESD emissions estimates were given preference over data from other sources.

Example of ESD Data Processing Steps
Source Category: Carbon Black Production

ESD Data Template Screening Sheet:
Section A:
Screeners were asked to review the templates and indicate on the Screening Sheet if there were facility-specific emissions data or nation-wide emissions data provided, and a preliminary description of the data provided in the template. If sufficient data were not provided to proceed with the processing, processing was stopped at this point, and the reason for stoppage documented.
Section B:
Record counts of each template, and the number of unique facilities and stacks listed in the templates were recorded. Based on these record counts, it was decided whether the templates should be processed at the facility-level or national-level, or not further processed.

ESD Data Processing Summary Sheet
If the templates for a source category made it past the initial screening, the four separate templates were transformed into a single linear format, and documentation for the 5 processing steps were recorded on the Processing Summary Sheet.
1. Duplicate Check between Existing NTI and ESD File: Data such as TRI ID codes, site name, latitude and longitude, SCC, SIC, state and county fips, and zip code were used to find duplicate emissions records. A list of what queries were used to find duplicates, or what facilities were manually found were listed in this section for QA purposes. Where duplicate facilities were found, two tables were created: ESD Dupes and NTI Dupes.
2. Steps Related to the Separation of Area Source Data from Major Source Data: Steps and relevant notes for the separation of area source data and major source data are listed here. Where area source data was available, a separate table in access was created for that data.
3. Application of Selection Criteria to Determine if ESD Data is Preferable to Existing NTI Data Where Duplicates are Found: Written documentation is provided for the application of the selection criteria. A selection criteria column was used in the ESD Dupes and NTI Dupes tables. From these tables, based on the selection criteria, three additional tables were created: ESD Point Adds (ESD data to be added to the major source inventory), ESD Area Adds (ESD data to be added to the area source inventory), and NTI Deletes (state data to be removed from the major source inventory and replaced by ESD data). The numeric codes for the selection criteria are provided in Table 1.
4. Steps to Supplement Either the Existing NTI Data with ESD Data or vice-versa: If ESD data were selected over existing NTI data, but the existing NTI data provided some additional information, such as SIC or zip code, the ESD data was supplemented with these data. Written documentation for the use of such supplemental data was recorded in section 4.
5. Other Processing Steps Performed: Any other manipulation of the data is documented in section 5.

ESD-to-ORACLE Data Template Fact Sheet
This is the complementary QA form to the ESD Data Processing Summary Sheet. It provides summary data for the tables created as a result of the ESD Data Processing. A separate Template

Fact Sheet is used for each table created. Information regarding the emission detail level; unique sites, units, and processes; and data integrity checks were documented. In this carbon black example, the number of unique facilities matches the number of unique facilities documented in the Template Screening Sheet. As was documented in the Processing Summary Sheet, no state data were given preference over the ESD data, and there was no area source data that needed to be separated from the major source data; so the number of unique facilities in the Template Screening Sheet and the Template Fact Sheet are the same. This will not necessarily be the case with each source category.

Processing QA Review Form
This documents that the proper files have been created at each step of the data processing, and are located in their respective directories. It also checks that all the required tables had been created as a result of the processing. It documents that each of the five steps listed in the data processing summary sheet have been completed properly.

Oracle Ready (Access) Tables: ESD Adds - Line Record Checks
After the Processing QA Review Form has been completed and the data has met the QA requirement outlined in that form, the data is then separated into the separate tables of the NTI format. This QA form documents that these separate NTI-format files recreate the data contained in the linear files of the data processing. Queries were used to compare site name, site id codes, SIC, location data, unit and process data, reporting year, stack data, emissions, and control equipment data.

ESD Data Template Screening Sheet

MACT Category:_____

MACT ID: _____

ESD File Path and Name: _____

ESD Data Screener: _____

A) Initial Screening of Data received from ESD

All screeners should review, if available, the "Attachment G: Summary of 1996 Templates Populated by ESD" that was returned by ESD. Attachment G may provide information to help perform the initial screening of the ESD data; however, in all cases screeners must review the actual ESD files themselves in the event there is contradiction in what the Attachment G states and what data actually appears in the files.

1) Are there emissions data in the Emissions Template ? _____

2) Are there emissions data in the National Template ? _____

3) Are the emissions in the National Template 100 % Area ? _____

4) What is the base year of the emissions ? _____

5) Are there location data (at the very least a state and city/county name) by facility ?_____

In the following table put "X" in the blank for each table for which ESD provided data. If additional fields were provided, then please provide their names and whether they should be included in the final ORACLE-Ready table.

Template Name	Data Received	Additional Field Names	Include Additional Field in ORACLE-Ready Table?
National Template			
Facility Template			
Stack Parameters Template			
Emission Template			

6) Do you believe based on the questions asked above and the initial indication of what was provided in the ESD templates that the ESD file(s) for this source category can be processed any further?   Yes   No

If "No", please describe below why you believe the ESD data files can not be processed further at this time:

1

ESD Data Template Screening Sheet

B) Initial Data Query Results: Part B should be completed if, based on the results of Part A above, you determine that the ESD data file can be processed further. At this point, you will need to copy the ESD template tables into Access in order to compile the data queries requested below. In the blanks provided, please put the results of the query. If the query is not applicable for the data provided (e.g.. the Stack Parameters Template is blank), then put "no data").

Query Number	Query Name	National Template	Facility Template	Stack Parameters Template	Emissions Template
1	QA: count of records - National Template		NA	NA	NA
2	QA: count of records - Facility Template	NA		NA	NA
3	QA: count of records – Stack Parameters Template	NA	NA		NA
4	QA: count of records - Emissions Template	NA	NA	NA	
5	QA: # major/area source facilities - Facility Template	NA	Area = Major =	NA	NA
6	QA: national M/A split		NA	NA	NA
7	QA: # unique facilities - Facility Template	NA		NA	NA
8	QA: # unique facilities - Stack Parameters Template	NA	NA		NA
9	QA: # unique facilities – Emissions Template	NA	NA	NA	
10	QA: # unique stacks – Stack Parameters Template	NA	NA		NA
11	QA: # unique stacks – Emissions Template	NA	NA	NA	

Do you believe based on the query results listed above and any further knowledge of what was provided in the ESD templates that the ESD file(s) for this source category can be processed any further?   Yes   No

If "Yes", what is the next step in processing required (indicate any special instructions for further processing that you feel are necessary for the data processor to follow) ?

1) the ESD data files require facility-level emissions data processing.

2) the ESD data file requires national-level emissions data processing.

If "No", please describe below why you believe the ESD data files can not be processed further at this time:

2

ESD Data Processing Summary Sheet

MACT Category: _____

MACT ID: _____

Name of Access File Received by Processor: _____

Name of Final Access File Created by Processor: _____

Data Processor: _____

The purpose of this form is to have written documentation of what data processing steps were performed on the ESD file. Enough information should be provided so that a reviewer can follow the processing steps from beginning to end. It is not necessary to document the results of all your queries and trivial data manipulations. The purpose is to document the critical steps and judgement calls that are made along the way. You MUST retain a copy of your final Access database that contains all the critical queries that you used along the way (e.g., the query and results to perform the duplicate check). That database will be backed-up on CD and stored as additional documentation of the processing steps.

Below are subheadings for critical points already identified in the data flow—at a minimum, processing steps performed at these critical points needs to be documented below each step. Other processing steps not directly related to these critical points, can be described under the "Other Processing Steps" heading.

1) **Duplicate Check Between Existing NTI and ESD File**

2) **Steps Related to the Separation of Area Source Data from Major Source Data**

3) **Application of Selection Criteria to Determine if ESD Data is Preferable to Existing NTI Data Where Duplicates are Found**

4) **Steps to Supplement Either the Existing NTI Data with ESD Data or visa-versa**

5) **Other Processing Steps Performed**

ESD-to-ORACLE Data Template Fact Sheet

MACT Category: _____

MACT ID: _____

Data Processor: _____

This fact sheet should be created for any ESD ADD or NTI REVISED table that is created as part of the ESD processing steps. The fact sheet helps to summarize the new and revised data that will be added to the Oracle database. The fact sheet does NOT replace the need for written documentation of the processing steps that were used to process the ESD data. That written documentation should be done on the "Data Processing Summary Sheet".

Provide all information requested for the following items. If an item is not applicable, indicate "NA" in the space next to the item.

A) Access File Information:

ORACLE-Ready Access File Name and Path: _____

ORACLE-Ready Access Table Name: _____

B) Level of Data Aggregation:
Circle the number corresponding to the level at which emissions data have been processed in the Oracle-ready Access Table identified above. Note: You may have different levels for different facilities within a source category—in these cases, you will have separate ORACLE-Ready Access tables corresponding to the level at which the data have been processed (e.g., you can have a table for facility level totals for certain facilities in the source category, and a table for process-level emissions for other facilities in the same category—in this example, you will create a fact sheet for both tables).

1	Facility totals only (i.e., there is only 1 emission value given for each pollutant at a facility and there is no stack, emission unit, process ID, or SCC given for the emission value.)
2	Stack totals only (i.e., there is only a single stack estimate and a single fugitive estimate for every facility/pollutant pair and there are no emission units, process IDs, or SCCs given for the emission value.)
3	Stack level (i.e., there exist at least 2 stacks with emissions data for a given facility/pollutant pair and there are no emission units, process IDs, or SCCs given for the emission value.)
4	Emission unit level (i.e., emissions are provided for each emission unit ID at a facility and there are no process IDs or SCCs given for the emission value.)
5	Process level (i.e., emissions are provided for each SCC, process description, or ESD defined process code)

C) Facility Check
1) Number of unique facilities: _____
2) Name of query that verifies uniqueness: _____
3) Facilities are uniquely defined by which of the following fields (circle one):
   site_ID , site_name , federal_id_code_2, other (write in field name here) _____
4) Was the field that uniquely defines a facility provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique facility): _____
5) Each unique facility must be linked with, at most, one SIC. Place an X to verify this was checked: _____

ESD-to-ORACLE Data Template Fact Sheet

D) Stack ID Check

1) Was the field that uniquely defines a Stack ID provided by ESD (if not, then list the fields from the ESD templates that were combined to create the unique emission unit): _____

2) Number of unique facility and stack ID pairs: _____

3) Name of query that verifies uniqueness: _____

E) Emission Unit Check

1) Number of unique emission unit IDs: _____

2) Name of query that verifies uniqueness: _____

3) Emission units are uniquely defined by grouping on the unique facility field circled above and (circle one): state_local_emis_unit_id , emis_unit_description

4) Was the field that uniquely defines an emission unit provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique emission unit): _____

F) SCC or Process Check

1) Number of unique SCCs: _____

2) Name of query that verifies uniqueness: _____

3) SCCs are uniquely defined by grouping on the unique facility field and the unique emission unit field circled above (if applicable) and (circle one): SCC , state_local_process_id , process_description

4) Was the field that uniquely defines an SCC provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique SCC): _____

5) Are the emissions given for a single or multiple time periods (e.g.., is it all 1996 emissions or are there some 1994, 1995, and 1996 emissions? _____

G) Other Checks:

1) Are the emissions given for a single or multiple time periods (i.e., is it all 1996 emissions or are there some 1994, 1995, and 1996 emissions? _____

2) How many different emission types are there? _____

3) What is the total number of populated emissions fields ? (e.g., 10 records with each having 2 emission types means you've got 20 emissions) _____

4) What is the total sum of all emissions (all types, all units and not converted to a common unit) ? _____

5) Are latitude and longitude site location data provided (yes/no): _____

6) Are latitude and longitude emission unit location data provided (yes/no): _____

7) Are latitude and longitude stack ID location data provided (yes/no): _____

8) Are stack parameter data provided (yes/no): _____

9) Is the control device type field populated (yes/no): _____

10) Is the control status field populated (yes/no): _____

ESD-to-ORACLE Data Template Fact Sheet

Notes: _____
_____
_____
_____
_____
_____
_____
_____
_____

PROCESSING QA REVIEW FORM
(See also attached flowchart)

MACT Category: _____

MACT Category: ID: _____

Name of Procesor's Access File: _____

QA Reviewer: _____

Area Access File Created by Reviewer: _____

Point Access File Created by Reviewer: _____

___ Verify that the completed database named on the ESD Data Processing Summary Sheet is located in:

N:[MACT\NTIESD96\PROCQA

___ Verify that the completed database has the following:

Original table
Add/Delete/Revised tables
Queries/Intermediate steps used in processing

___ Check original table with Stephanie's linear table to verify that they are the same. The linear table can be found at:

N:\MACT\NTIESD96\STEPH\LINEAR

___ Flag all judgmental calls made by processor for Joe Mangino to review.

___ Verify that a column called Cut Criteria was added to all DUPES tables and that data for this column has been properly entered.

___ Check the five standard steps on the ESD Data Processing Summary Sheet.
1. ___
2. ___
3. ___
4. ___
5. ___

___ If problems are found that are not simple, contact the processor and tell them what the problems are so that they can be fixed.

If returned to processor:

List problems found: _____

_____

_____

Date returned to processor: _____

1

Date received back from processor: _____

Amended documentation provided: _____

___ Either re-start or continue with this checklist (depending on what was re-done by processor). Re-start ____ Continue ____

___ Fill out the ESD-to-ORACLE Data Template Fact Sheet.

___ When you have determined that the tables are final, copy final tables to new databases as instructed in the next check point. Retain the Full original database in the PROCQA folder. The final tables can consist of:

ESD POINT ADD
ESD AREA ADD
NTI DELETE

One or the other of these two tables will be created:
NTI REVISED (make sure this one has the MACT_ID field completed)
NTI POINT REVISED (make sure this one has the MACT_ID field completed)

The following two tables (which should be identical) represent data that will be moved out of point source NTI into the area source NTI:
NTI AREA ADD
NTI POINT DELETE

___ If there are ESD AREA ADD and/or NTI AREA ADD tables, you need to create a new database that contains only the Area tables with the filenaming convention of:

XXXXXFNA.MDB (Ex: LEATHFNA.MDB)

Then move this file to:

N:\MACT\NTIESD96\AREAFIN

___ Create a new database that contains only the Point Source related tables (i.e., ESD POINT ADD, NTI DELETE, NTI REVISED or NTI POINT REVISED) with the filenaming convention of:
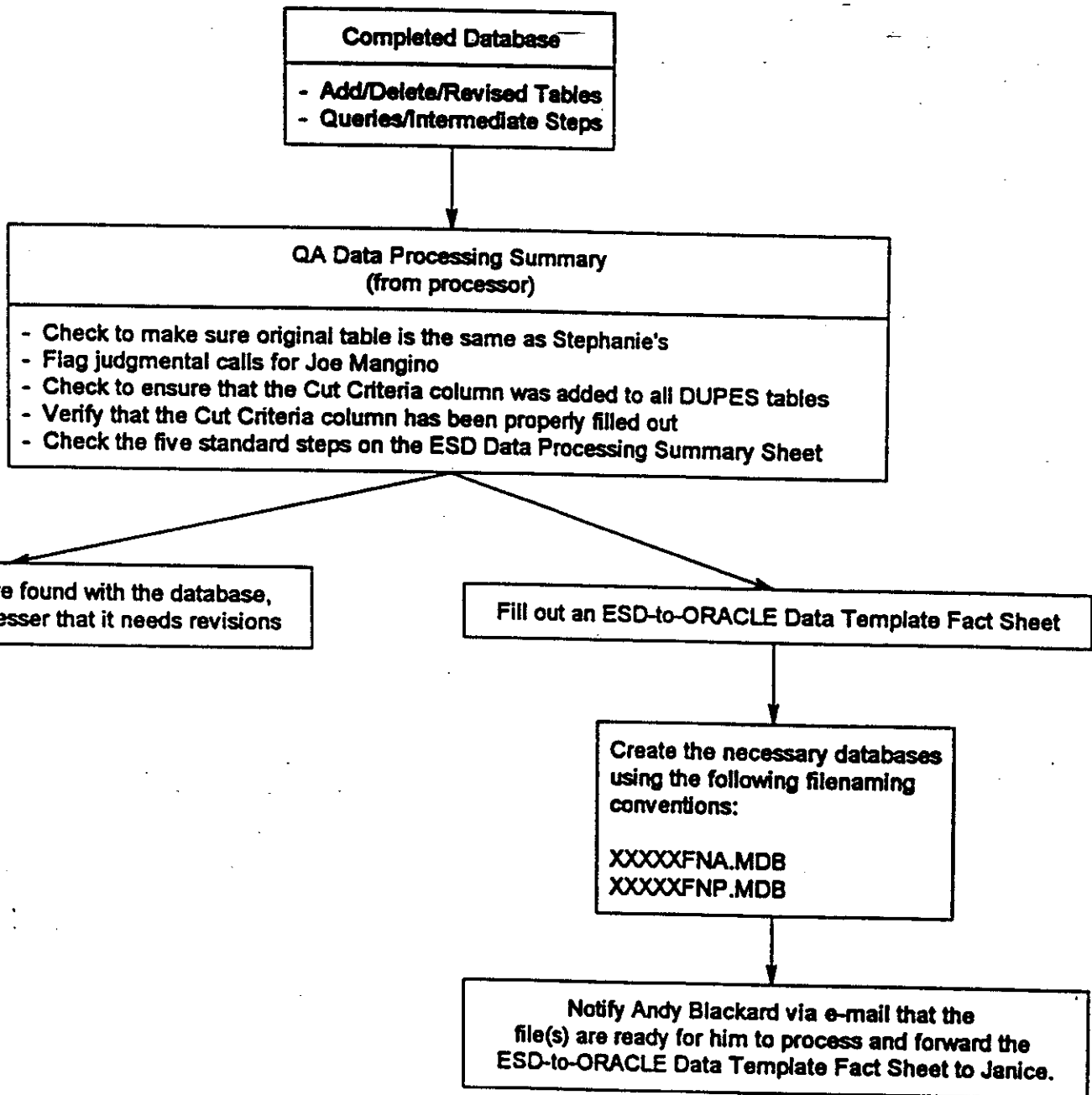
XXXXXFNP.MDB (Ex: LEATHFNP.MDB)

Then move this file to:

N:\MACT\NTIESD96\POINTFIN

___ Notify Andy Blackard by e-mail that the table(s) are ready for him to process. Give him the name(s) and location(s) of the file(s) and forward the ESD-to-ORACLE Data Template Fact Sheet to Janice at that time.

2

PROCESSING QA REVIEW FLOWCHART

```
                    ┌─────────────────────────────────┐
                    │       Completed Database         │
                    │                                  │
                    │  - Add/Delete/Revised Tables     │
                    │  - Queries/Intermediate Steps    │
                    └─────────────────────────────────┘
                                    │
                                    ▼
┌──────────────────────────────────────────────────────────────────┐
│               QA Data Processing Summary                           │
│                   (from processor)                                 │
├────────────────────────────────────────────────────────────────── │
│  - Check to make sure original table is the same as Stephanie's    │
│  - Flag judgmental calls for Joe Mangino                           │
│  - Check to ensure that the Cut Criteria column was added to all DUPES tables │
│  - Verify that the Cut Criteria column has been properly filled out│
│  - Check the five standard steps on the ESD Data Processing Summary Sheet │
└──────────────────────────────────────────────────────────────────┘
```

If problems are found with the database, notify the processer that it needs revisions	Fill out an ESD-to-ORACLE Data Template Fact Sheet

```
                                    │
                                    ▼
                    ┌─────────────────────────────────┐
                    │  Create the necessary databases  │
                    │  using the following filenaming  │
                    │  conventions:                    │
                    │                                  │
                    │  XXXXXFNA.MDB                    │
                    │  XXXXXFNP.MDB                    │
                    └─────────────────────────────────┘
                                    │
                                    ▼
            ┌───────────────────────────────────────────────┐
            │  Notify Andy Blackard via e-mail that the      │
            │  file(s) are ready for him to process and forward the │
            │  ESD-to-ORACLE Data Template Fact Sheet to Janice. │
            └───────────────────────────────────────────────┘
```

QA CHECKS:
ORACLE READY (ACCESS) TABLES
ESD ADDS - LINE RECORD CHECKS

Please run these checks on the final 14 tables in the databases in the n:\....\template\janice directory. The Import Me table can be obtained from the n:\...\template directory.

Mact Category:_____

Mact ID:_____

Number of Records in Import Me Table:_____

Number of Records in Linearized Query:_____

Are all the fields with data listed in the queries?_____

Reviewer:_____

Results of Queries:

Z01 Compare Site Name: _____

Z02 Compare Federal Ids:_____

Z03 Compare SICs:_____

Z04 Compare Locations:_____

Z05 Compare Units & Processes:_____

Z06 Compare Times:_____

Z07 Compare Stacks:_____

Z08 Compare Emissions:_____

Z09 Compare Controls:_____

NOTES:_____

_____

_____

ESD Data Template Screening Sheet

MACT Category: _____ CARBON BLACK PRODUCTION

MACT ID: _____ 1415

ESD File Path and Name: _____ N/Mact/Ntiesd96/Esdrecv/2_15_99/Nti-212/Ocg/Cblack.mdb

ESD Data Screener: _____ AHL

A) Initial Screening of Data received from ESD

All screeners should review, if available, the "Attachment G: Summary of 1996 Templates Populated by ESD" that was returned by ESD. Attachment G may provide information to help perform the initial screening of the ESD data; however, in all cases screeners must review the actual ESD files themselves in the event there is contradiction in what the Attachment G states and what data actually appears in the files.

1) Are there emissions data in the Emissions Template ? _____ YES

2) Are there emissions data in the National Template ? _____ YES

3) Are the emissions in the National Template 100 % Area ? _____ NO

4) What is the base year of the emissions ? _____ 1997

5) Are there location data (at the very least a state and city/county name) by facility ? _____ YES

In the following table put "X" in the blank for each table for which ESD provided data. If additional fields were provided, then please provide their names and whether they should be included in the final ORACLE-Ready table.

Template Name	Data Received	Additional Field Names	Include Additional Field in ORACLE-Ready Table?
National Template	X	NA	
Facility Template	X	NA	
Stack Parameters Template	X	NA	
Emission Template	X	NA	

6) Do you believe based on the questions asked above and the initial indication of what was provided in the ESD templates that the ESD file(s) for this source category can be processed any further?   Yes   No   YES

If "No", please describe below why you believe the ESD data files can not be processed further at this time:

1

ESD Data Template Screening Sheet

B) Initial Data Query Results: Part B should be completed if, based on the results of Part A above, you determine that the ESD data file can be processed further. At this point, you will need to copy the ESD template tables into Access in order to compile the data queries requested below. In the blanks provided, please put the results of the query. If the query is not applicable for the data provided (e.g., the Stack Parameters Template is blank), then put "no data").

Query Number	Query Name	National Template	Facility Template	Stack Parameters Template	Emissions Template
1	QA: count of records - National Template	3	NA	NA	NA
2	QA: count of records - Facility Template	NA	20	NA	NA
3	QA: count of records -- Stack Parameters Template	NA	NA		NA
4	QA: count of records - Emissions Template	NA	NA	NA	51
5	QA: # major/area source facilities - Facility Template	NA	Area = Major =20	NA	NA
6	QA: national M/A split		NA	NA	NA
7	QA: # unique facilities - Facility Template	NA	20	NA	NA
8	QA: # unique facilities - Stack Parameters Template	NA	NA	20	NA
9	QA: # unique facilities – Emissions Template	NA	NA	NA	20
10	QA: # unique stacks – Stack Parameters Template	NA	NA	95	NA
11	QA: # unique stacks – Emissions Template	NA	NA	NA	20

Do you believe based on the query results listed above and any further knowledge of what was provided in the ESD templates that the ESD file(s) for this source category can be processed any further?   Yes   No   **YES**

If "Yes", what is the next step in processing required (indicate any special instructions for further processing that you feel are necessary for the data processor to follow) ?

1) the ESD data files require facility-level emissions data processing.

CONTINUE WITH FACILITY LEVEL EMISSIONS PROCESSING. NOTE THAT THE EMISSIONS TEMPLATE DOES NOT CONTAIN STACK IDS SO UNLESS THE ESD CONTACT CAN PROVIDE STACK IDS, AVERAGE STACK PARAMETER VALUES WILL NEED TO BE CALCULATED AND USED WHEN CREATING THE LINEAR TABLE.

2) the ESD data file requires national-level emissions data processing.

If "No", please describe below why you believe the ESD data files can not be processed further at this time:

ESD Data Processing Summary Sheet

MACT Category: <u>Carbon Black Production</u>

MACT ID: <u>1415</u>

Name of Access File Received by Processor: <u>n:\mact\ntiesd96\steph\linear\cblack1/mdb</u>

Name of Final Access File Created by Processor: <u>cblack2.MDB</u>

Data Processor: <u>Michael Bryant</u>

1) Duplicate Check Between Existing NTI and ESD File

As described below, I found duplicate for eleven (11) of the twenty (20) facilities, and this data is contained in the ESD POINT ADD table,:

[A] Cabot (Three Facilities): Canal, Ohio River, & Ville Platte.
[B] Columbian Chemicals (Four Facilities): El Dorado, Franklin, Moundsville, & Ulysses.
[C] Continental Carbon (Two Facilities): Phenix City, & Ponca City
[D] Degaussa/Belpre,
[E] Sid Richardson/Addis.

2) Steps Related to the Separation of Area Source Data from Major Source Data

The ESD data file contained no data from area sources. As described below, the table ESD ADD contains records for the nine (9) facilities that are not duplicates.

[A] Cabot/Pampa
[B] Continental Carbon/Sunray
[C] Degaussa (Two Facilities): Aransas Pass, & Ivanhoe
[D] Engineered Carbon (Three Facilities): Baytown, Borger, & Echo.
[E] Sid Richardson (Two Facilities): Big Springs, & Borger.

The NTI DELETE table contains the data for the eleven duplicate facilities

[A] Cabot (Three Facilities): Canal, Ohio River, & Ville Platte.
[B] Columbian Chemicals (Four Facilities): El Dorado, Franklin, Moundsville, & Ulysses.
[C] Continental Carbon (Two Facilities): Phenix City, & Ponca City
[D] Degaussa/Belpre
[E] Sid Richardson/Addis

3) **Application of Selection Criteria to Determine if ESD Data is Preferable to Existing NTI Data Where Duplicates are Found**

The original NTI data was excluded using criteria 3 & 4 for the following six (6) facilities:

[A] **Cabot/Ohio River**
[B] **Columbian Chemicals (Two Facilities): El Dorado, & Moundsville**
[C] **Continental Carbon (Two Facilities): Phenix City, & Ponca City**
[D] **Degaussa/Belpre**

The remaining original NTI data was excluded using criteria 7 for the following five (5) facilities:

[A] **Cabot (Two Facilities): Canal, & Ville Platte.**
[B] **Columbian Chemicals (Two Facilities): Centerville, & Ulysses.**
[C] **Sid Richardson/Addis**

4) **Steps to Supplement Either the Existing NTI Data with ESD Data or visa-versa**

The ESD POINT ADD table contains additional data in the SIC_NEW column.

5) **Other Processing Steps Performed**

Created three tables: (1) ESD ADD, (2) NTI DELETE, and (3) ESD POINT ADD

ESD-to-ORACLE Data Template Fact Sheet ✓

MACT Category: _Carbon Black Production_

MACT ID: _1415_

Data Processor: _Michael Bryant_

QA: Colleen Kane 3/26/99

This fact sheet should be created for any ESD ADD ~~or MFI-REVISED~~ table that is created as part of the ESD processing steps. The fact sheet helps to summarize the new and revised data that will be added to the Oracle database. The fact sheet does NOT replace the need for written documentation of the processing steps that were used to process the ESD data. That written documentation should be done on the "Data Processing Summary Sheet".

Provide all information requested for the following items. If an item is not applicable, indicate "NA" in the space next to the item.

A) Access File Information:

ORACLE-Ready Access File Name and Path: _n:\mact\rtic add\point fin\cblacfng.mdb_

ORACLE-Ready Access Table Name: _ESD Point Add_

B) Level of Data Aggregation:
Circle the number corresponding to the level at which emissions data have been processed in the Oracle-ready Access Table identified above. Note: You may have different levels for different facilities within a source category—in these cases, you will have separate ORACLE-Ready Access tables corresponding to the level at which the data have been processed (e.g., you can have a table for facility level totals for certain facilities in the source category, and a table for process-level emissions for other facilities in the same category—in this example, you will create a fact sheet for both tables).

1	Facility totals only (i.e., there is only 1 emission value given for each pollutant at a facility and there is no stack, emission unit, process ID, or SCC given for the emission value.)
2	Stack totals only (i.e., there is only a single stack estimate and a single fugitive estimate for every facility/pollutant pair and there are no emission units, process IDs, or SCCs given for the emission value.)
3	Stack level (i.e., there exist at least 2 stacks with emissions data for a given facility/pollutant pair and there are no emission units, process IDs, or SCCs given for the emission value.)
4	Emission unit level (i.e., emissions are provided for each emission unit ID at a facility and there are no process IDs or SCCs given for the emission value.)
(5)	Process level (i.e., emissions are provided for each SCC, process description, or ESD defined process code)

C) Facility Check
1) Number of unique facilities: _20_
2) Name of query that verifies uniqueness: _esd point add unique facil check_
3) Facilities are uniquely defined by which of the following fields (circle one):
   site_ID , site_name , federal_id_code_2, other (write in field name here) _state_local_site_ID_code_

   ~~Was the field that uniquely defines a facility provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique facility):_____~~
5) Each unique facility must be linked with, at most, one SIC. Place an X to verify this was checked: _X_

D) Stack ID Check

X Was the field that uniquely defines a Stack ID provided by ESD (if not, then list the fields from the ESD templates that were combined to create the unique emission unit): _____

2) Number of unique facility and stack ID pairs: _20_

3) Name of query that verifies uniqueness: _esd point add unique stack id check_

E) Emission Unit Check

1) Number of unique emission unit IDs: _20_

2) Name of query that verifies uniqueness: _esd point add unique emission unit check_

3) Emission units are uniquely defined by grouping on the unique facility field circled above and (circle one): (state_local_emis_unit_id) emis_unit_description

X Was the field that uniquely defines an emission unit provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique emission unit): _____

F) SCC or Process Check

1) Number of unique SCCs: _20_

2) Name of query that verifies uniqueness: _esd point add unique scc check_

3) SCCs are uniquely defined by grouping on the unique facility field and the unique emission unit field circled above (if applicable) and (circle one): (SCC) state_local_process_id , process_description

X Was the field that uniquely defines an SCC provided by ESD (if not, then list the fields from the ESD templates that were combined to identify a unique SCC): _____

5) Are the emissions given for a single or multiple time periods (e.g... is it all 1996 emissions or are there some 1994, 1995, and 1996 emissions? _single_   _1997_

G) Other Checks:

1) Are the emissions given for a single or multiple time periods (i.e., is it all 1996 emissions or are there some 1994, 1995, and 1996 emissions? _Single_   _1997_

2) How many different emission types are there? _1_

3) What is the total number of populated emissions fields ? (e.g., 10 records with each having 2 emission types means you've got 20 emissions) _48_

4) What is the total sum of all emissions (all types, all units and not converted to a common unit) ? _5682_

5) Are latitude and longitude site location data provided (yes/no): _yes_

6) Are latitude and longitude emission unit location data provided (yes/no): _no_

7) Are latitude and longitude stack ID location data provided (yes/no): _no_

8) Are stack parameter data provided (yes/no): _no_

9) Is the control device type field populated (yes/no): _yes_

10) Is the control status field populated (yes/no): _no_

Notes: _____

_____

_____

_____

_____

_____

_____

_____

_____

PROCESSING QA REVIEW FORM
(See also attached flowchart)

MACT Category: _Carbon Black Production_
MACT Category: ID: _1415_
Name of Procesor's Access File: _cblack2.mdb_
QA Reviewer: _Colleen Kane_
Area Access File Created by Reviewer: _n/a_
Point Access File Created by Reviewer: _cblacfnp.mdb_

✓   Verify that the completed database named on the ESD Data Processing Summary Sheet is located in:

     N:[MACT\NTIESD96\PROCQA

✓   Verify that the completed database has the following: _✓checked for Maine already_

                                                    dupes — 11 facil

     ✓ Original table _20 facil, 48 records_
     ✓ Add/Delete/Revised tables
     ✓ Queries/Intermediate steps used in processing

✓   Check original table with Stephanie's linear table to verify that they are the same. The linear table can be found at:

     N:\MACT\NTIESD96\STEPH\LINEAR

N/A   Flag all judgmental calls made by processor for Joe Mangino to review.

✓   Verify that a column called <u>Cut Criteria</u> was added to all DUPES tables and that data for this column has been properly entered.

✓   Check the five standard steps on the ESD Data Processing Summary Sheet. _I combined Miller ESD add & ESD Point ADD into one table_
     1.   ✓   _• I moved sic_new into sic for dupe records in ESD Point add_
     2.   ✓   _& deleted sic_new field._
     3.   ✓
     4.   ✓   _→ 2 new tables ESD Point add_
     5.   ✓   _+ NTI Delete_

N/A   If problems are found that are not simple, contact the processor and tell them what the problems are so that they can be fixed.

If returned to processor:

List problems found: _____

_____

_____

Date returned to processor: _____

Date received back from processor: _____

Amended documentation provided: _____

___ Either re-start or continue with this checklist (depending on what was re-done by processor). Re-start ____ Continue ____

✓ Fill out the ESD-to-ORACLE Data Template Fact Sheet

✓ When you have determined that the tables are final, copy final tables to new databases as instructed in the next check point. Retain the Full original database in the PROCQA folder. The final tables can consist of:

✓ ESD POINT ADD
✗ ESD AREA ADD
✓ NTI DELETE

One or the other of these two tables will be created:
✗ NTI REVISED (make sure this one has the MACT_ID field completed)
✗ NTI POINT REVISED (make sure this one has the MACT_ID field completed)

The following two tables (which should be identical) represent data that will be moved out of point source NTI into the area source NTI:
NTI AREA ADD
✗ NTI POINT DELETE

N/A If there are ESD AREA ADD ~~and/or NTI AREA ADD tables~~, you need to create a new database that contains only the Area tables with the filenaming convention of:

XXXXXFNA.MDB (Ex: LEATHFNA.MDB)

Then move this file to:

N:\MACT\NTIESD96\AREAFIN

✓ Create a new database that contains only the Point Source related tables (i.e., ESD POINT ADD, NTI DELETE, NTI REVISED or NTI POINT REVISED) with the filenaming convention of:

XXXXXFNP.MDB (Ex: LEATHFNP.MDB)
cblacfnp.mdb
Then move this file to:

N:\MACT\NTIESD96\POINTFIN

✓ Notify Andy Blackard by e-mail that the table(s) are ready for him to process. Give him the name(s) and location(s) of the file(s) and forward the ESD-to-ORACLE Data Template Fact Sheet to him at that time.

QA CHECKS:
ORACLE READY (ACCESS) TABLES
ESD ADDS - LINE RECORD CHECKS

Please run these checks on the final 14 tables in the databases in the n:\....\template\janice directory. The Import Me table can be obtained from the n:\...\template directory.

Mact Category: _Carbon Black_

Mact ID: _1415_

Number of Records in Import Me Table: _48_

Number of Records in Linearized Query: _48_

Are all the fields with data listed in the queries? _yes_

Reviewer: _Steph Fenn_

Results of Queries:

Z01 Compare Site Name: _ok_

Z02 Compare Federal Ids: _ok   state local ID is really fed ID 2_

Z03 Compare SICs: _OK_

Z04 Compare Locations: _OK_

Z05 Compare Units & Processes: _emission unit type descr is missing_

Z06 Compare Times: _OK_

Z07 Compare Stacks: _OK_

Z08 Compare Emissions: _OK_

Z09 Compare Controls: _ok_

NOTES: _emission unit type # should be 06, not 17_
_also: state_local_emiss_unit_ID was_
truncated; in order to match Input me
did a temp fix - do we need to address
this?